



Gene expression profiling of mouse host response to *Listeria monocytogenes* infection

Hanna H. Ng*, Christopher E. Frantz, Linda Rausch, David C. Fairchild, Julie Shimon, Edward Riccio, Sherrod Smith, Jon C. Mirsalis

Biosciences Division, SRI International, 333 Ravenswood Avenue, Menlo Park, CA 94025-3493, USA

Received 3 May 2005; accepted 14 July 2005

Available online 15 August 2005

Abstract

The purpose of this study was to evaluate gene expression profiles in the liver and blood for prediction of infection severity from *Listeria monocytogenes* (LM). Mice were injected with medium broth (control) or a nonlethal or lethal dose of LM and sacrificed 6 h later. Gene expression changes were determined using Affymetrix MGU74Av2 GeneChips and confirmed by real-time polymerase chain reaction analysis. We identified discernable genes whose gene expression profiles can be used in pattern recognition to predict and classify samples in differently treated groups, with $\geq 90\%$ accuracy in liver samples and 80% accuracy in blood at prediction; however, different genes were predictive in each tissue. Our results suggest that gene expression profiling in response to LM in mice may be able to distinguish samples in groups with varying severity of infection and provide information in finding molecular mechanisms and early biomarkers for subsequent conventional clinical endpoints.

© 2005 Elsevier Inc. All rights reserved.

Keywords: *Listeria monocytogenes*; Gene expression profiling; Microarray; Class prediction; Biomarker

There is a tremendous need for research studies on pathogen-induced characteristics in host gene expression, with the ultimate goal being the design and development of assays that can rapidly identify and diagnose pathogen exposure before the onset of overt clinical symptoms. We chose *Listeria monocytogenes* (LM) as a model infectious agent because it is a food-borne pathogen that affects pregnant women, newborns, and immunocompromised adults. Listeriosis has been recognized as an important public health problem in the United States [1]. The systemic infection of mice with LM is also a widely used model to study mammalian host defense response to intracellular bacterial pathogens [2] and to examine immunotoxic effects of therapeutics [3].

LM is an intracellular gram-positive pathogen [2]. Immediately after systemic infection, most bacteria are

cleared rapidly from the bloodstream and can be recovered in the liver 10 min after infection [4]. The vast majority of microorganisms are entrapped in the liver and destroyed there by phagocytes and neutrophils [2,4]. The listerial burden in the liver increases exponentially after 6 h, reaching a plateau between days 3 and 4 and then declining with an increase in specific immunity [5]. The LM bacterium utilizes the host's cellular processes to affect cell-to-cell interactions, move intracellularly, and proliferate [6]. Much of this is accomplished through pathogen-induced alterations in host cell signal transduction [7] and by utilizing the host cell's actin-based cytoskeleton [6]. While LM is a well-studied model, the effects of LM on host gene expression remain largely unknown. Characterization of the LM-induced alterations in mouse gene expression will provide mechanistic information on the disease process and yield biomarkers of clinical disease.

Advances in microarray technology have enabled massive parallel mining of gene expression data by measuring the hybridization of mRNA to thousands of specific cDNA

* Corresponding author. Fax: +1 (650) 859 3444.

E-mail address: hanna.ng@sri.com (H.H. Ng).

pieces or oligonucleotides immobilized on the surface of a “chip” that are complementary to known genes or expressed sequence tags (ESTs). Microarrays provide a sensitive and high-throughput tool that can be applied to many areas, including cancer research, molecular and cell biology, physiology, pharmacology, toxicology, and risk assessment [8–12]. Using microarrays, multiple genes have been identified that have altered gene expression in response to chemical exposure in cell culture [13] and in animals [14–16]. Gene transcription profiles have been proposed to be useful in discriminating samples exposed to different classes of toxicants and predicting toxicity of unknown or similar compounds [17]. In addition, analyzing gene expression patterns across the tumor tissues of individual patients demonstrated the potential of refinement or confirmation in cancer classification that provides better prognostication and treatment selection [18–24]. Thus, expression profiling may classify and predict phenotypes of toxicity in response to bioagents or chemical exposures.

Infection by microorganisms leads to a series of changes in the physiology of host cells. Many changes are relayed by signal transduction systems and are expressed by alterations in the transcription levels of host genes. Numerous studies using microarray technology have shown that infectious microorganisms, including human immunodeficiency virus 1 [25], simian immunodeficiency virus [26], herpes simplex virus 1 [27], adenovirus [28], *Coxiella burnetii* or *Chlamydia trachomatis* [29], *Mycobacterium tuberculosis* [30,31], *Salmonella* [32], and the protozoan kinetoplastid parasite *Trypanosoma cruzi* [33], induce differential gene expression in host target cells during the course of infection. However, most reported microarray studies have centered on primary cells or cultured cell lines. Few examples [26,31,33] have described the host response to systemic infection with virus, bacteria, and protozoans in vivo. To our knowledge this is the first report of microarray gene expression studies in host response to LM.

Since gene expression responses to infectious agents begin to take place within the first few hours after exposure, gene expression profiling will be useful in diagnosing pathogen exposure in advance of clinical symptoms, allowing exposed individuals to be rapidly triaged for treatment. Our goal is to characterize the changes in gene expression in female BALB/c mice 6 h after infection with LM and evaluate the gene expression profiling in the liver and blood for prediction of clinical outcomes.

Results

Infectious burden, histopathology, and clinical pathology evaluation

Female BALB/c mice were administered, via intravenous tail injection, medium broth (control) or 9.8×10^2 colony

forming units (CFU) (nonlethal dose) or 5.3×10^4 CFU (lethal dose) of LM. The LM replicated rapidly in the liver, reaching a plateau between days 1 and 2 (Fig. 1). Then the infectious burden in the mice infected with the nonlethal dose gradually declined, clearing by day 21. LM burden data are not available for the lethal group on days 4, 14, and 21 because there were no survivors in this group after day 3. There was no detectable LM in the livers from control (uninfected) animals.

Pathological evaluations of brain and liver that were harvested at various times postinfection were performed on animals receiving control broth, nonlethal dose, and lethal dose of LM. The brains from all mice were histologically normal. Ten percent of mice that received the nonlethal dose and 100% mice that received the lethal dose had hepatic inflammatory lesions (abscessation) characteristic of LM infection (Fig. 2). In addition, significant increases in serum aspartate aminotransferase and alanine aminotransferase, which are markers of liver toxicity, were observed in the lethal dose group (Table 1).

LM-induced gene expression changes in the host

Liver samples were processed for microarray gene expression analysis; 31 GeneChips (9 from control animals, 13 from the nonlethal group, and 9 from the lethal group) passed the quality checks after hybridization and scanning. After normalization and filtering for gene quality control, 6227 genes were left for further evaluation. Since the minimum change in differential expression that can be accurately detected is 1.4-fold [34], we selected only those genes for which differential expression changes were 1.5-fold or greater and were statistically significant (one-way ANOVA, $p < 0.05$). We found 53 genes in the nonlethal group and 89 genes in the lethal group that had at least a 1.5-fold change in transcription level, compared with those in the control group; 17 genes were significantly changed in both the nonlethal and the lethal group, compared with controls. These differentially expressed genes in the LM-

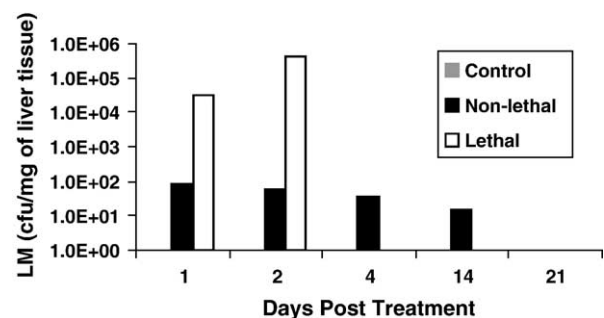


Fig. 1. Recovery of LM from the liver after injection with medium broth (control), nonlethal dose, and lethal dose. There was no detectable LM in the livers from control animals. Animals in the lethal high-dose group were found dead on day 3 or were severely moribund and therefore sacrificed. No data were obtained for the lethal group on days 4, 14, and 21. $N = 3$ animals per sex at each time point.

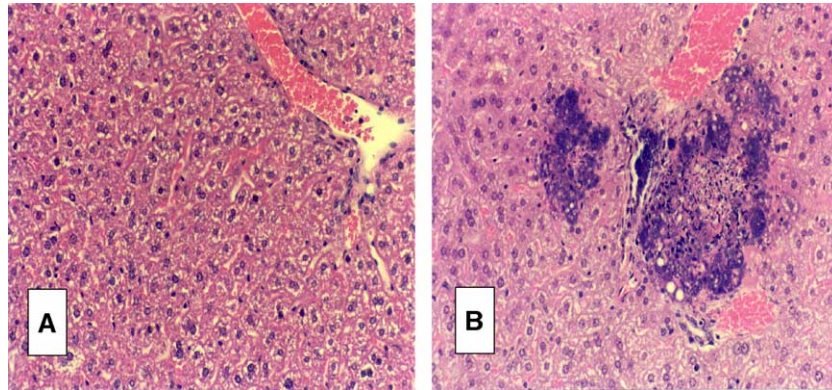


Fig. 2. Histopathology of panel A normal control liver tissue and panel B liver tissue from mouse with lethal infection with LM. In (A), the vein present in one corner of the control is partially filled with erythrocytes. In (B), there are two foci of necrotic hepatic parenchyma containing myriad bacteria. The parenchyma surrounding lesions is normal. Original magnification 100 \times ; hematoxylin and eosin stain.

infected animals are to be further studied to provide mechanistic information on the disease process and mammalian host defense response to intracellular bacterial pathogens.

Sample prediction using LM-induced host gene expression in liver

We sought to identify genes with a prediction rule to classify samples into groups correlating with clinical outcomes. Employing the GeneSpring software, we used an approach similar to that of Golub et al. [20]. In our study, 21 liver gene expression profiles, 7 from each group (control, nonlethal, lethal), were randomly selected in the Training Set to build the model using the *K*-nearest-neighbor algorithm to predict and classify the rest of 10 independent unknown samples according to their gene expression profiles. Fifty genes were selected arbitrarily as highly ranked predictive genes in the class prediction model. In the prediction, 9 of the 10 test samples were correctly predicted and 1 was incorrect in terms of infection severity. Similar predictions with differently randomized samples in the Training Set and Test Set were performed two more times. Three sets of 50 arbitrary predictive genes were

generated. A small subset of 8 genes was identified to be common in these three sets of 50 predictive genes. We used the gene expression signature of this 8-gene set in all 31 samples to construct a hierarchical clustering according to the expression similarities between samples. As shown in Fig. 3, gene expression patterns were significantly associated with the treatment groups; animals in the same group as control, nonlethal, or lethal clustered together.

We then used this subset of eight genes to predict outcomes in the 10 Test samples, which randomized three times from the 31 liver samples. As shown in Table 2, of 10 unknown Test samples, 9 or all 10 samples were correctly predicted. The overall prediction accuracy was $\geq 90\%$. The sample (No. 30) that was predicted incorrectly was an animal in the nonlethal group predicted to be a control sample. This was consistent with the sample clustering data (in Fig. 3, sample “30_liver nonlethal”), indicating that this animal may have been less responsive to the infection.

The eight predictive genes included four upregulated and four downregulated genes and were characterized with regard to gene ontology classes according to the query from the Affymetrix database as shown in Table 3. The fold change of the gene represented a relative transcriptional level in the treated animals compared with that in controls.

Real-time RT-PCR for confirmation of expression change

Real-time RT-PCR was used to confirm the gene expression fold changes generated by using microarray analysis. Among the eight predictive genes in liver, we chose two upregulated genes, T-cell-specific GTPase (*Tgtp*) and serum amyloid A3 gene (*Saa3*), and one downregulated gene, cytochrome P450, 17 (*Cyp17a1*). An internal control (*Gapdh*, a housekeeping gene) was used as a baseline to analyze the relative levels of *Tgtp*, *Saa3*, and *Cyp17a1* in RT-PCR. As shown in Table 3, the RT-PCR results support the microarray-based observation of the fold change in gene

Table 1

Clinical chemistry parameters: aspartate aminotransferase (AST) and alanine aminotransferase (ALT)

Group	AST (IU/L)				ALT (IU/L)			
	Day 1	Day 2	Day 4	Day 14	Day 1	Day 2	Day 4	Day 14
Control	132	118	98	210	72	83	52	98
Nonlethal	199	138	152*	164	162*	115	117	67
Lethal	522*	8705*	–	–	339*	4609*	–	–

Blood was collected for clinical pathology from animals in each group on different days postinfection. The number is an average of $N = 6$, except in the lethal dose group on day 2, for which $N = 4$. One-way ANOVA followed by Dunnett's test was used to analyze differences between control and nonlethal or lethal samples.

* $p < 0.05$ was considered statistically significant.

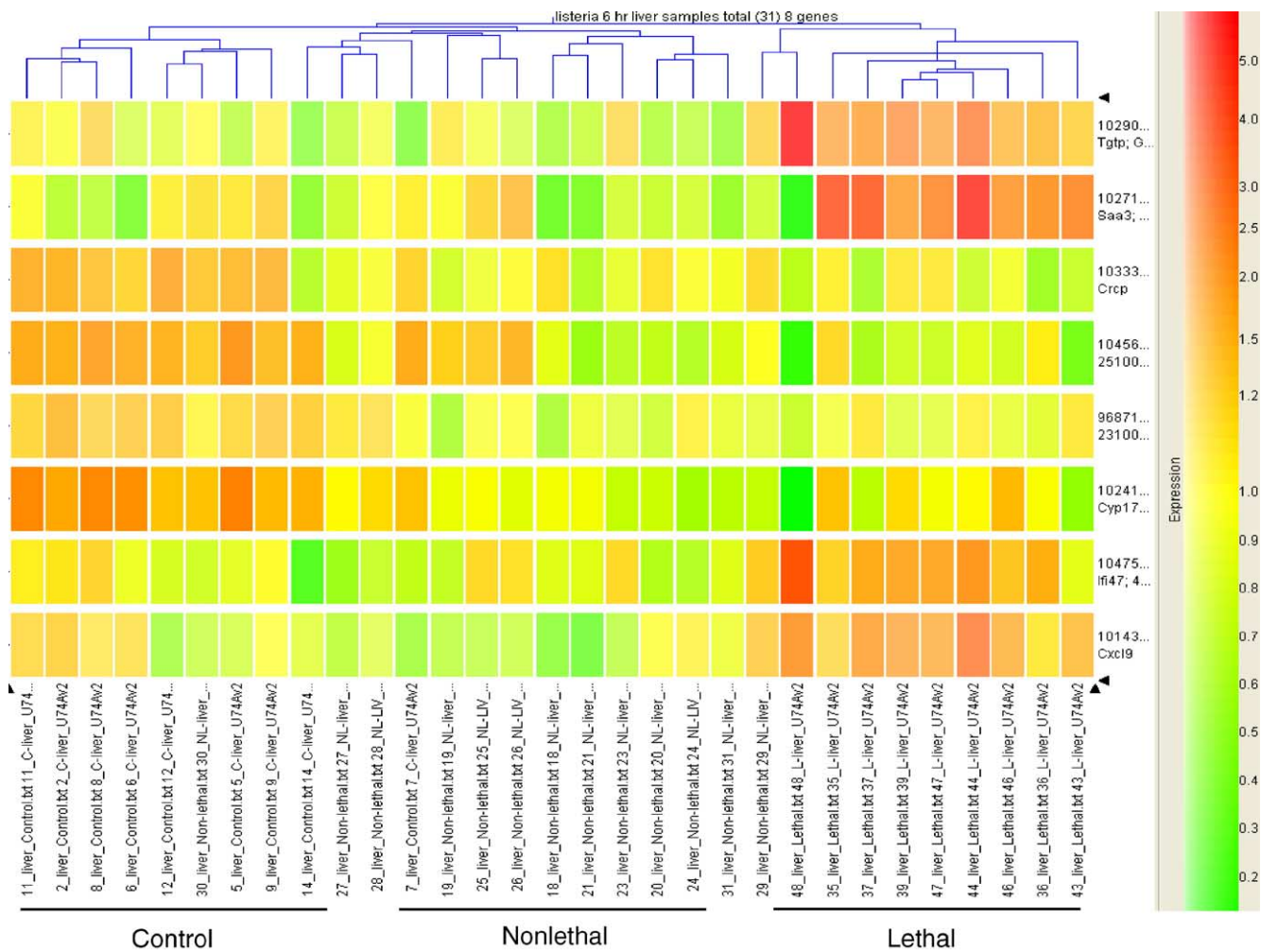


Fig. 3. Hierarchical clustering according to the similarities of the gene expression patterns for eight predictive genes between samples. Each row represents one of the eight gene expressions and each column a separate sample. Animals 1–14 were in the control group, 18–31 were in the nonlethal group, and 35–48 were in the lethal group. A pseudo-colored representation of relative intensity is shown such that red indicates high expression and green low expression, with scale shown at the right.

expression level in the nonlethal and lethal dose-infected animals relative to control animals.

Sample prediction using host gene expression profile in blood in response to LM

We applied similar class prediction methods in the analysis of 40 GeneChips that passed the initial quality check from blood samples collected from the animals treated with control, nonlethal, and lethal doses of LM as follows: 14 from control, 13 from nonlethal, and 13 from the lethal group. To build the model using the *K*-nearest-neighbor algorithm to predict and classify unknown samples, 30 gene expression profiles, 10 from each group, were randomly selected in the Training Set. Fourteen probes/genes were identified as high-ranking predictive genes, in that 2 probes on the GeneChip represent the same gene, *Bcl2a1a* (Table 4). We performed three rounds of randomization using 40 Gene-

Chips to assign 10 samples to the Test set. In each prediction, 8 of 10 unknown Test samples were correctly predicted (Table 5). The incorrectly predicted samples were animals in the nonlethal dose group that were predicted to be controls, suggesting less response to LM infection in these animals.

Discussion

The analysis of the interaction of pathogenic bacteria with their host cell on the molecular and cellular levels has become a major research area in recent years. Microarrays provide the advantage of being able to investigate the expression of thousands of genes, and bioinformatics software offers an effective way of managing complex, high-volume data and assisting in identifying patterns of gene expression changes [35]. When we used hierarchical cluster analysis to identify groups of samples related by

Table 2

Class prediction in three sets of randomized test samples from liver using the gene expression profiling for eight genes

Test 1			Test 2			Test 3		
Sample	True value	Prediction	Sample	True value	Prediction	Sample	True value	Prediction
12_liver	Control	Control	18_liver	Nonlethal	Nonlethal	18_liver	Nonlethal	Nonlethal
14_liver	Control	Control	21_liver	Nonlethal	Nonlethal	19_liver	Nonlethal	Nonlethal
19_liver	Nonlethal	Nonlethal	23_liver	Nonlethal	Nonlethal	20_liver	Nonlethal	Nonlethal
27_liver	Nonlethal	Nonlethal	25_liver	Nonlethal	Nonlethal	24_liver	Nonlethal	Nonlethal
28_liver	Nonlethal	Nonlethal	28_liver	Nonlethal	Nonlethal	29_liver	Nonlethal	Nonlethal
29_liver	Nonlethal	Nonlethal	30_liver	Nonlethal	Control ^a	2_liver	Control	Control
30_liver	Nonlethal	Control ^a	36_liver	Lethal	Lethal	31_liver	Nonlethal	Nonlethal
31_liver	Nonlethal	Nonlethal	43_liver	Lethal	Lethal	35_liver	Lethal	Lethal
47_liver	Lethal	Lethal	5_liver	Control	Control	44_liver	Lethal	Lethal
48_liver	Lethal	Lethal	8_liver	Control	Control	9_liver	Control	Control

The true value refers to the known exposure, control, nonlethal, or lethal. The prediction value refers to the class to which the animal was predicted to belong, based on the predictive gene set.

^a Incorrect prediction.

similarity in the liver gene expression profiles, the samples grouped in close proximity according to the control, nonlethal, or lethal infected animals (Fig. 3). Data from

this analysis demonstrated that gene expression in the host response is remarkably homogeneous within the same group but distinct from other groups. One outlier was animal 30,

Table 3

Characterization of subset of eight predictive genes from liver with regard to gene ontology^a classes and relative gene expression level as determined by microarray and RT-PCR analysis

Probe set ID (Affymetrix)	Gene name	Gene symbol	Biological process description	Molecular function description	Cellular component description	Gene expression fold change (microarray) ^b		Gene expression fold change (RT-PCR) ^c	
						Non-lethal	Lethal	Non-lethal	Lethal
103334_at	calcitonin gene-related peptide-receptor component protein	<i>Crcp</i>	Neuropeptide signaling pathway		Acrosome	0.69	0.58	–	–
101436_at	chemokine (C-X-C motif) ligand 9	<i>Cxcl9</i>	Immune response Inflammatory response	Cytokine activity Chemokine activity	Extracellular space Extracellular	0.81	2.01	–	–
102906_at	T-cell-specific GTPase	<i>Tgtp</i>		GTP binding		1.04	3.43	1.40	3.31
102416_at	cytochrome P450, family 17, subfamily a, polypeptide 1	<i>Cyp17a1</i>	Electron transport C21-steroid hormone biosynthesis	Monooxygenase activity Steroid 17- α -monooxygenase activity Oxidoreductase activity	Mitochondrion Membrane	0.54	0.53	0.63	0.45
104750_at	Interferon- γ -inducible protein	<i>Ifi47</i>	Defense response			1.00	1.97	–	–
104566_at	argininosuccinate lyase	<i>Asl</i>	Arginine catabolism Arginine biosynthesis Urea cycle	Lyase activity Catalytic activity Argininosuccinate lyase activity	Cytoplasm	0.58	0.44	–	–
96871_at	RIKEN cDNA 2310042G06 gene	<i>2310042G06Rik</i>				0.69	0.71	–	–
102712_at	serum amyloid A3	<i>Saa3</i>	Acute-phase response	Lipid transporter activity	Extracellular space	1.00	3.25	1.94	3.08

^a Gene ontology information was derived from the Affymetrix NetAffx Analysis Center.

^b Microarray data: fold change of gene expression level in treated groups relative to the control group value, which was normalized to 1. A number >1 means the gene is upregulated, and a number <1 means the gene is downregulated. $N = 9–13$ animals/group.

^c RT-PCR analysis was performed for three genes: fold change of gene expression level in treated groups is relative to the control group value. $N = 3–4$ animals/group.

Table 4

Characterization of subset of 14 predictive genes in blood with regard to gene ontology^a classes and relative gene expression level in microarray

Probe set ID (Affymetrix)	Gene name	Gene symbol	Biological process description	Molecular function description	Cellular component description	Gene expression fold change (microarray) ^b	
						Non-lethal	Lethal
100311_f_at	eosinophil-associated, ribonuclease A family, member 1	<i>Ear1</i>		Nucleic acid binding, nuclease activity, endonuclease activity, pancreatic ribonuclease activity, hydrolase activity	Extracellular space	2.61	5.25
100569_at	annexin A2	<i>Anxa2</i>	Angiogenesis, collagen fibril organization, fibrinolysis	Phospholipase inhibitor activity, calcium ion binding, protein binding, calcium-dependent phospholipid binding, cytoskeletal protein binding	Stress fiber, membrane fraction, cytoplasm, early endosome, cell junction	1.38	3.4
101571_g_at	insulin-like growth factor binding protein 4	<i>Igfbp4</i>	Regulation of cell growth	Insulin-like growth factor binding, growth factor binding	Extracellular, extracellular space	3.47	3.23
102091_f_at						1.6	1.63
102914_s_at	B-cell leukemia/lymphoma 2-related protein A1a, A1b, A1c, A1d	<i>Bcl2a1a</i> , <i>Bcl2a1b</i> , <i>Bcl2a1c</i> , <i>Bcl2a1d</i>	Apoptosis or regulation of apoptosis			1.91	2.51
103089_at	CD48 antigen	<i>Cd48</i>			Extracellular space, plasma membrane	1.45	1.49
103240_f_at	eosinophil-associated, ribonuclease A family, member 2	<i>Ear2</i>	Chemotaxis	Nucleic acid binding, nuclease activity, endonuclease activity, pancreatic ribonuclease activity, hydrolase activity		1.66	2.65
103830_at	snail homolog 1 (<i>Drosophila</i>)	<i>Snail</i>	Development, organogenesis	Nucleic acid binding, DNA binding, zinc ion binding	Nucleus	0.66	0.57
104482_at	epimorphin	<i>Epim</i>	Intracellular protein transport	Protein transporter activity	Membrane, integral to membrane	0.84	0.72
162142_f_at	proline-rich Gla (G-carboxyglutamic acid) polypeptide 2	<i>Prrg2</i>		Calcium ion binding, protein binding	Extracellular region, extracellular space, integral to membrane	0.65	0.79
93869_s_at	B-cell leukemia/lymphoma 2-related protein A1a, A1c, A1d	<i>Bcl2a1a</i> , <i>Bcl2a1c</i> , <i>Bcl2a1d</i>	Apoptosis or regulation of apoptosis			1.29	2.08
94061_at	cysteine-rich protein 1 (intestinal)	<i>Crip1</i>		Zinc ion binding		1.01	2.58
96522_at	calmodulin 1	<i>Calm1</i>	Cell cycle, G-protein-coupled receptor protein signaling pathway	Calcium ion binding, protein binding	Cytoplasm, plasma membrane	1.42	1.79
98092_at	placenta-specific 8	<i>Plac8</i>				1.55	2.5

^a Gene ontology information was derived from the Affymetrix NetAffx Analysis Center.^b Fold change of gene expression level in treated groups relative to the control group value, which was normalized to 1. A number >1 means the gene is upregulated, and a number <1 means the gene is downregulated. *N* = 13–14 animals/group.

which was infected with a nonlethal dose of LM but clustered with the control samples, indicating that this may have been a less responsive animal. It also correlates with the data in Table 2 in class prediction analysis that animal 30 was predicted as a control animal.

Fifty or as few as 8 genes with high-ranking scores according to their significance for classification in the *K*-

nearest-neighbor algorithm were identified to be able to predict the animals treated with varying doses of LM. In the 8-predictive gene set, 4 genes, including serum amyloid A3 (*Saa3*), T-cell-specific GTPase (*Tgtp*), an interferon (IFN)- γ -inducible protein (*Ifi47*), and chemokine (C-X-C motif) ligand 9 (*Cxcl9*), were upregulated. SAA3 is an acute-phase response protein and is upregulated by a

Table 5

Class prediction in three sets of randomized test samples using a subset of 14 predictive genes in blood

Test 1			Test 2			Test 3		
Sample	True value	Prediction	Sample	True value	Prediction	Sample	True value	Prediction
21_blood	Nonlethal	Control ^a	23_blood	Nonlethal	Not predicted ^b	11_blood	Control	Control
25_blood	Nonlethal	Nonlethal	28_blood	Nonlethal	Nonlethal	12_blood	Control	Control
28_blood	Nonlethal	Nonlethal	31_blood	Nonlethal	Control ^a	18_blood	Nonlethal	Not predicted ^b
36_blood	Lethal	Lethal	42_blood	Lethal	Lethal	24_blood	Nonlethal	Nonlethal
37_blood	Lethal	Lethal	45_blood	Lethal	Lethal	29_blood	Nonlethal	Control ^a
39_blood	Lethal	Not predicted ^b	47_blood	Lethal	Lethal	36_blood	Lethal	Lethal
4_blood	Control	Control	4_blood	Control	Control	37_blood	Lethal	Lethal
6_blood	Control	Control	5_blood	Control	Control	38_blood	Lethal	Lethal
7_blood	Control	Control	8_blood	Control	Control	5_blood	Control	Control
8_blood	Control	Control	9_blood	Control	Control	9_blood	Control	Control

The true value refers to the known exposure, control, nonlethal, or lethal. The prediction value refers to the class to which the animal was predicted to belong, based on the predictive gene set.

^a Incorrect prediction.

^b Not predicted because *p* value ratio >0.2.

variety of inflammatory stimuli, including cytokines and glucocorticoids [36,37]. SAA stimulates the rapid expression and release of tumor necrosis factor- α from cultured human blood neutrophils and monocytes, suggesting a role for SAA in cytokine production in acute inflammation and immune responses [38]. *Tgtp* is a gene encoding a 47-kDa protein in mouse peritoneal macrophages induced by IFN- γ [39]. This gene was strongly induced by endogenous IFN- $\alpha\beta$, suggesting that it might participate in defense against viral infection [40]. Studies indicated that IFN- γ and interleukin-12 are key cytokines that induce type 1 immune responses in protection of the host against LM infection [2]. C-X-C chemokines are major chemoattractants that direct granulocytes, natural killer cells, macrophages, and lymphocytes migrating into inflammatory sites [41]. Many chemokines are induced by IFN- γ [42]. The upregulation of predictor genes *Saa3*, *Tgtp*, *Ifi47*, and *Cxcl9* is consistent with the functions of these genes in the host immune or inflammatory process in response to viral and bacterial invasions. The downregulation of the other predictor genes, including an unknown RIKEN DNA sequence, is less obvious, although it was found that hepatic cytochrome p450 enzymes CYP1A, CYP2B, and CYP3A were significantly lowered 48 and 72 h after systemic infection with LM [43], and this may be due to compromised hepatic function.

To validate the gene expression fold changes in microarray analysis, we employed a quantitative RT-PCR technique using hybridization probes in PCR, which gives specificity in measurement for the target amplification products and eliminates background noise. Not only did our RT-PCR results confirm that *Tgtp* and *Saa3* were upregulated and cytochrome P450 (*Cyp17a1*) was downregulated in the infected animals, but also the fold changes in expression levels were generally consistent. The RT-PCR showed a 1.4- and a 1.94-fold increase in *Tgtp* and *Saa3*, respectively, in the nonlethal group, which was not apparent in microarray analysis. This may be attributable to the mean

value of different numbers of animals being used in the two analyses and/or the quantitative RT-PCR using hybridization probes in detection format being more sensitive and specific at low levels and for small changes in mRNA levels.

Development of clinically relevant and sensitive surrogate markers of responses to the infectious agents prior to overt clinical toxicity is important. Our results indicated that using 14 genes from blood samples allows assignment of unknown samples to the control, nonlethal, or lethal group of LM infection with about 80% accuracy and thus may provide hope in finding surrogates for prediction of clinical outcomes; however, different genes were predictive in liver and blood. Two of the blood predictor genes are ESTs or without known function. The eosinophil-associated ribonuclease 1 (*Ear1*) gene was shown to play a role in host defense against RNA virus pathogens [44]. Annexin A2 is a member of the multifunctional annexin family of Ca^{2+} -dependent phospholipid binding proteins and has been reported to be upregulated in virally transformed cell lines [45]. The gene encoding insulin-like growth factor binding protein 4 (*Igfbp4*) is involved in regulation of tumor cell growth and was shown to be differentially expressed in astrocytoma progression [46]. This gene was also identified as one of the predictor genes that can differentiate hepatoblastoma tumor from nondiseased liver according to their gene expression profiles [47]. CD48 is a cell surface protein and a critical mediator of activation of mast cells following exposure to mycobacteria [48]. Other predictor genes may be involved in signal transduction, cell cycle regulation, and apoptosis (such as *Bcl2a1a*); however, the biological functions of these genes with regard to host response to LM infection are not clear. The relative importance and specific functions of the genes in the clearance of LM infection treated with antibiotics are currently under investigation in our laboratory.

We do not think, at this stage, that these predictor genes are specific for LM infection. There may also be differences in the host response gene expression between mouse,

which is permissive for listeriosis, and human, in which infection with LM is relatively innocuous in healthy people, but causes more severe effects in immunocompromised individuals. The predictor genes in the mouse model will therefore not necessarily be predictive in humans; however, given the fact that these genes represent broad response genes for stress and inflammation, it seems reasonable to assume that the same or similar genes will be relevant in human listeriosis. Further studies with a larger sample size may help to validate the predictive gene expression profiles and explore fully the entire network of the genes involved in the mechanism of response to LM infection. We acknowledge the fact that for some compounds or bioagents, the host response to the treatment may be governed by mechanisms that are not readily revealed at the transcriptional level (gene expression), such as post-transcriptional regulation, posttranslational modification, proteasome function, or protein-protein interactions. We are also aware that there is discordance between different array platforms, and standardization requires cooperation across industry, academia, and medicine, as well as across international borders [49].

Nevertheless, the present study shows that microarray analysis of host gene expression in response to pathogen infection provides a powerful approach to explore the expression of a large set of genes, which is important to understand the interaction between a pathogen and the host and pathogenesis. In addition, microarrays provide a tremendous amount of information that helps find expression patterns from different levels of exposure of LM and predict/classify unknown samples. Gene expression profiling in host blood using microarray and rapid PCR-based gene expression profiling in validation may be able to serve as a surrogate for toxicity outcome. This knowledge may ultimately contribute to the design of treatments for infectious disease, therapeutic development, and vaccination strategies.

Materials and methods

Animal treatment and sample collection

Female BALB/c mice (Charles River Laboratories) approximately 6–8 weeks of age were maintained on Purina Certified Rodent Chow 5002 (Richmond, IN, USA) and purified tap water ad libitum in microisolator cages under controlled lighting (12-h light-dark cycle). Animals were randomly assigned to dose groups and then administered, via intravenous tail injection, medium broth (control) or 9.8×10^2 to 2×10^3 CFU (nonlethal dose) or 5.3×10^4 to 7×10^4 CFU (lethal dose) of LM. Blood was collected via retro-orbital sinus and stored in PAXgene (Qiagen, Valencia, CA, USA) until RNA extraction. A section of liver was collected to quantify CFU of LM in liver for determination of pathogen burden, and a section of

liver was collected in phosphate-buffered 10% formalin for histopathology. The remaining portions of liver were stored in RNAlater (Qiagen) at -70°C until processed for RNA extraction.

Infection burden determination

At necropsy, the left lobe of the liver was weighed and placed in a 5-ml homogenizing tube kept on ice. Saline (1 ml) was added to the tube and the tissue was homogenized using a Polytron homogenizer. The homogenate was diluted as necessary and a 100- μl aliquot was plated (in duplicate) onto trypticase soy agar plates with 5% sheep blood (Becton–Dickinson, Sparks, MD, USA). The plates were incubated at 37°C for ~ 24 h, then bacterial plaques were counted and the infectious burden calculated as CFU/mg liver (average of four to six animals).

Histopathological analysis

The formalin-fixed liver samples were embedded in paraffin, cut approximately 5 μm thick, and stained with hematoxylin and eosin. Histopathologic examination was performed by a board-certified veterinary pathologist.

RNA isolation

Total RNA was extracted using the RNeasy kit (Qiagen). RNA quality was assessed by ethidium bromide agarose gel electrophoresis. Concentration was determined at 260 nm absorbance by a spectrophotometer.

Sample preparation for microarray hybridization and analysis

Total RNA was used to prepare labeled cRNA target for analysis on microarray according to the standard protocols from Affymetrix. Briefly, cDNA was synthesized from total RNA in the presence of T7-(dT)₂₄ oligomer, using the Superscript kit (Invitrogen, Carlsbad, CA, USA). The cRNA was prepared using biotin-labeled nucleotides in the ENZO Bioarray HighYield RNA Transcript Labeling Kit (Enzo Diagnostics, Farmingdale, NY, USA), then purified, quantified using spectrophotometric analysis, and checked by gel electrophoresis. The cRNA was fragmented using a fragmentation buffer and heat. The labeled cRNA was then added to a hybridization solution, hybridized to Murine Genome U74Av2 Arrays (MG-U74Av2) (Affymetrix, Santa Clara, CA, USA), and incubated at 45°C for 16 h. Following the hybridization step, the hybridized probe was washed and stained and then scanned to acquire an image of each probe. The image was analyzed for expression intensity using Affymetrix analysis software, Microarray Suite 5.0, and GeneSpring software (Silicon Genetics, Redwood City, CA, USA). Quality checks of the microarray screening included confirmation of complete cDNA synthesis of housekeeping

genes, positive detection of spiked control, and low and consistent background noise levels.

Microarray data analysis

Analysis of the data to find “interesting” genes involves first preprocessing the raw data for quality control, including standardization of the data (i.e., normalization to make the data uniformly comparable throughout the dataset), and filtering the data to eliminate genes with low precision, large error values, or expression levels that are too close to the background, based on the behaviors of the genes in experiments.

Experimental normalization

Global normalization was performed to standardize samples by minimizing nonbiological processing variation so that relative gene expression levels could be compared. We assumed that most of the genes were not changed across the GeneChips. Normalization was performed according to the suggested default procedures in the GeneSpring software.

Gene filters for quality control

We applied filters to retain only qualified genes that were “Present” or “Marginal” in the GeneChips. In blood sample analysis, since only about 10–15% of the total 12,422 mouse genes in the GeneChip were “Present” in the blood, we selected only genes with expression levels between treatment groups (control, nonlethal, and lethal) that were statistically significant (one-way ANOVA, $p \leq 0.05$).

Class prediction

We used a supervised learning method (K -nearest neighbor) in which the algorithm learns from samples with known class membership (Training Set) and establishes a prediction rule to classify independent new samples (Test Set). First, the transcription profiles of the animals in the control, nonlethal, or lethal group were used to identify genes with high predictive power for discerning groups by identifying idealized expression patterns associated with a given class. Following the identification, a “neighborhood analysis” was performed, in which the number of genes correlated to an idealized expression pattern was counted and the results were compared with a randomly generated pattern. A p value was calculated for the likelihood of observed representation of each class. The class with the smallest p value was compared to the class with the second smallest p value (p value ratio), and a prediction was made if the p value ratio was <0.2 .

Hierarchical clustering

A centroid-clustering method was used to calculate the profiles representing a binary subtree of samples in the similarity search. The algorithm first found the two most

correlated gene expression profiles and grouped them under one subtree. Then, it found the pair of the subtree that was the next most correlated and grouped that pair under one branch. This process was repeated until all the genes were merged into one large tree.

Validation of predictive gene expression by real-time RT-PCR

Real-time RT-PCR was performed in a two-step process. In the first step, the single stranded cDNA was synthesized from total RNA using SuperScript II RNase H⁻ reverse transcriptase (Invitrogen). In the second step, quantitative measurement of target transcript in each animal sample was achieved using a LightCycler with FastStart DNA Master Hybridization Probes (Roche, Indianapolis, IN, USA). The use of hybridization probes in PCR gives specificity in measurement for the target amplification products and eliminates background noise. The thermal cycling conditions for PCR were initial denaturation for 10 min at 95°C, followed by 50 cycles of 95°C for 15 s, annealing at 55°C for 10 s, extension at 72°C for 15 s, and fluorescent data acquisition at 55°C. A standard curve for each primer pair was created in each PCR run by using four to seven serially diluted reference cDNA samples from a control animal. Gene-specific primers for real-time PCR corresponding to the genes present on the Affymetrix GeneChip were designed and synthesized (TIB MolBio LLC, Adelphia, NJ, USA). The following primers were used: mouse T-cell-specific GTPase (mTGTP), forward 5'-GAAGTGT-TATTGCCACCAGATC-3' and reverse 5'-GCACTCTC-GATGTCTCTCAGTA; mouse serum amyloid A3 protein (SAA3), forward 5'-AAGAAGCTGGTCAAGGGTCT and reverse 5'-CTCTCTGGCATCGCTGATG; cytochrome P450, 17 (CYP17), forward 5'-CGGTTTATGCCTGAGCG and reverse 5'-CGATCAGAAAGACCACCTTG; and mouse GAPDH, forward 5'-ATTCAACGGCACAGT-CAAGG and reverse 5'-TGGATGCAGGGATGATGTTC-3'. GAPDH was amplified by using the same target cDNA as a control to normalize the expression levels of the genes. In preparation of cDNA, the total RNA was from the same animal but in different extractions, and therefore, they were parallel but independent processes. For RT-PCR analysis, duplicate PCRs were carried out for each cDNA sample, and the cDNA of three to four animals from each treatment group was used in the RT-PCR confirmation.

Acknowledgments

We thank Charles Litterst at the National Institute of Allergy and Infectious Diseases, NIH, for his scientific contributions and review of the manuscript, Sandra Phillips at SRI for her assistance for the in-life portion of the study, and Elizabeth Zuo and Al Smith at the Stanford University Protein and Nucleic Acid Facility for

the hybridization and image scan of Affymetrix microarray chips. This work was supported by NIAID Contract N01-AI-05417.

References

- [1] L. Slutsker, A. Schuchat, in: E.T. Ryser, E.H. Marth (Eds.), *Listeria, Listeriosis, and Food Safety*, Dekker/Little, Brown, New York, 1999, pp. 75–95.
- [2] S.H. Kaufmann, Immunity to intracellular bacteria, *Annu. Rev. Immunol.* 11 (1993) 129–163.
- [3] Y.R. Freund, E.S. Riccio, S.J. Phillips, L. Dousman, J.T. MacGregor, Pyrimethamine impairs host resistance to infection with *Listeria monocytogenes* in BALB/c mice, *Toxicol. Sci.* 42 (1998) 91–98.
- [4] S.H. Gregory, A.J. Sagnimeni, E.J. Wing, Bacteria in the bloodstream are trapped in the liver and killed by immigrating neutrophils, *J. Immunol.* 157 (1996) 2514–2520.
- [5] E.J. Wing, A. Waheed, R.K. Shaddock, Changes in serum colony-stimulating factor and monocytic progenitor cells during *Listeria monocytogenes* infection in mice, *Infect. Immun.* 45 (1984) 180–184.
- [6] T. Chakraborty, Molecular and cell biological aspects of infection by *Listeria monocytogenes*, *Immunobiology* 201 (1999) 155–163.
- [7] M. Kuhn, T. Pfeuffer, L. Greiffenberg, W. Goebel, Host cell signal transduction during *Listeria monocytogenes* infection, *Arch. Biochem. Biophys.* 372 (1999) 166–172.
- [8] M.L. Cunningham, M.S. Bogdanffy, T.R. Zacharewski, R.N. Hines, Use of genomic data in risk assessment, *Toxicol. Sci.* 73 (2003) 209–215.
- [9] J.E. Staunton, D.K. Slonim, H.A. Collier, P. Tamayo, M.J. Angelo, J. Park, et al., Chemosensitivity prediction by transcriptional profiling, *Proc. Natl. Acad. Sci. USA* 98 (2001) 10787–10792.
- [10] C.A. Afshari, E.F. Nuwaysir, J.C. Barrett, Application of complementary DNA microarray technology to carcinogen identification, toxicology, and drug safety evaluation, *Cancer Res.* 59 (1999) 4759–4760.
- [11] W.D. Pennie, J.D. Tugwood, G.J. Oliver, I. Kimber, The principles and practice of toxigenomics: applications and opportunities, *Toxicol. Sci.* 54 (2000) 277–283.
- [12] H.K. Hamadeh, S. Jayadev, E.T. Gaillard, Q. Huang, R. Stoll, K. Blanchard, et al., Integration of clinical and gene expression endpoints to explore furan-mediated hepatotoxicity, *Mutat. Res.* 549 (2004) 169–183.
- [13] F. de Longueville, F.A. Atienzar, L. Marcq, S. Dufrane, S. Evrard, L. Wouters, et al., Use of a low-density microarray for studying gene expression patterns induced by hepatotoxicants on primary cultures of rat hepatocytes, *Toxicol. Sci.* 75 (2003) 378–392.
- [14] Q. Huang, R.T. Dunn 2nd, S. Jayadev, O. DiSorbo, F.D. Pack, S.B. Farr, et al., Assessment of cisplatin-induced nephrotoxicity by microarray technology, *Toxicol. Sci.* 63 (2001) 196–207.
- [15] H.K. Hamadeh, P.R. Bushel, S. Jayadev, K. Martin, O. DiSorbo, S. Sieber, et al., Gene expression analysis reveals chemical-specific profiles, *Toxicol. Sci.* 67 (2002) 219–231.
- [16] J.A. Kramer, S.W. Curtiss, K.L. Kolaja, C.L. Alden, E.A. Blomme, W.C. Curtiss, et al., Acute molecular markers of rodent hepatic carcinogenesis identified by transcription profiling, *Chem. Res. Toxicol.* 17 (2004) 463–470.
- [17] H.K. Hamadeh, P.R. Bushel, S. Jayadev, O. DiSorbo, L. Bennett, L. Li, et al., Prediction of compound signature using high density gene expression profiling, *Toxicol. Sci.* 67 (2002) 232–240.
- [18] S.L. Pomeroy, P. Tamayo, M. Gaasenbeek, L.M. Sturla, M. Angelo, M.E. McLaughlin, et al., Prediction of central nervous system embryonal tumour outcome based on gene expression, *Nature* 415 (2002) 436–442.
- [19] M. Bittner, P. Meltzer, Y. Chen, Y. Jiang, E. Seftor, M. Hendrix, et al., Molecular classification of cutaneous malignant melanoma by gene expression profiling, *Nature* 406 (2000) 536–540.
- [20] T.R. Golub, D.K. Slonim, P. Tamayo, C. Huard, M. Gaasenbeek, J.P. Mesirov, et al., Molecular classification of cancer: class discovery and class prediction by gene expression monitoring, *Science* 286 (1999) 531–537.
- [21] C.H. Yeang, S. Ramaswamy, P. Tamayo, S. Mukherjee, R.M. Rifkin, M. Angelo, et al., Molecular classification of multiple tumor types, *Bioinformatics* 17 (Suppl. 1) (2001) S316–S322.
- [22] J. Khan, J.S. Wei, M. Ringner, L.H. Saal, M. Ladanyi, F. Westermann, et al., Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks, *Nat. Med.* 7 (2001) 673–679.
- [23] J. DeRisi, L. Penland, P.O. Brown, M.L. Bittner, P.S. Meltzer, M. Ray, et al., Use of a cDNA microarray to analyse gene expression patterns in human cancer, *Nat. Genet.* 14 (1996) 457–460.
- [24] I. Hedenfalk, D. Duggan, Y. Chen, M. Radmacher, M. Bittner, R. Simon, et al., Gene-expression profiles in hereditary breast cancer, *N. Engl. J. Med.* 344 (2001) 539–548.
- [25] G.K. Geiss, R.E. Bumgarner, M.C. An, M.B. Agy, A.B. van 't Wout, E. Hammersmark, et al., Large-scale monitoring of host cell gene expression during HIV-1 infection using cDNA microarrays, *Virology* 266 (2000) 8–16.
- [26] M.D. George, S. Sankaran, E. Reay, A.C. Gelli, S. Dandekar, High-throughput gene expression profiling indicates dysregulation of intestinal cell cycle mediators and growth factors during primary simian immunodeficiency virus infection, *Virology* 312 (2003) 84–94.
- [27] N.N. Khodarev, S.J. Advani, N. Gupta, B. Roizman, R.R. Weichselbaum, Accumulation of specific RNAs encoding transcriptional factors and stress response proteins against a background of severe depletion of cellular RNAs in cells infected with herpes simplex virus 1, *Proc. Natl. Acad. Sci. USA* 96 (1999) 12062–12067.
- [28] H. Zhao, F. Granberg, L. Elfneh, U. Pettersson, C. Svensson, Strategic attack on host cell gene expression during adenovirus infection, *J. Virol.* 77 (2003) 11006–11015.
- [29] Q. Ren, S.J. Robertson, D. Howe, L.F. Barrows, R.A. Heinzen, Comparative DNA microarray analysis of host cell transcriptional responses to infection by *Coxiella burnetii* or *Chlamydia trachomatis*, *Ann. N. Y. Acad. Sci.* 990 (2003) 701–713.
- [30] Y. Xu, J. Xie, Y. Li, J. Yue, J. Chen, L. Chunyu, et al., Using a cDNA microarray to study cellular gene expression altered by *Mycobacterium tuberculosis*, *Chin. Med. J. (Engl.)* 116 (2003) 1070–1073.
- [31] P. Seiler, P. Aichele, S. Bandermann, A.E. Hauser, B. Lu, N.P. Gerard, et al., Early granuloma formation after aerosol *Mycobacterium tuberculosis* infection is regulated by neutrophils via CXCR3-signaling chemokines, *Eur. J. Immunol.* 33 (2003) 2676–2686.
- [32] L. Eckmann, J.R. Smith, M.P. Housley, M.B. Dwinell, M.F. Kagnoff, Analysis by high density cDNA arrays of altered gene expression in human intestinal epithelial cells in response to infection with the invasive enteric bacteria *Salmonella*, *J. Biol. Chem.* 275 (2000) 14084–14094.
- [33] S. Mukherjee, T.J. Belbin, D.C. Spray, D.A. Jacobas, L.M. Weiss, R.N. Kitsis, et al., Microarray analysis of changes in gene expression in a murine model of chronic chagasic cardiomyopathy, *Parasitol. Res.* 9 (2003) 187–196.
- [34] H. Yue, P.S. Eastman, B.B. Wang, J. Minor, M.H. Doctolero, R.L. Nuttall, et al., An evaluation of the performance of cDNA microarrays for detecting changes in global mRNA expression, *Nucleic Acids Res.* 29 (2001) E41.
- [35] A. Grewal, J. Stockton, C. Bolger, Tools for discovery: gene expression enterprise solutions, *Curr. Opin. Drug Discovery Dev.* 6 (2003) 333–338.
- [36] R.S. Stearman, C.A. Lowell, W.R. Pearson, J.F. Morrow, Regulation of synthesis of amyloid A-related protein, *Ann. N. Y. Acad. Sci.* 389 (1982) 106–115.
- [37] C.F. Thorn, Z.Y. Lu, A.S. Whitehead, Regulation of the human acute phase serum amyloid A genes by tumour necrosis factor- α , interleukin-6 and glucocorticoids in hepatic and epithelial cell lines, *Scand. J. Immunol.* 59 (2004) 152–158.

- [38] E. Hatanaka, C.J. Furlaneto, F.P. Ribeiro, G.M. Souza, A. Campa, Serum amyloid A-induced mRNA expression and release of tumor necrosis factor- α (TNF- α) in human neutrophils, *Immunol. Lett.* 91 (2004) 33–37.
- [39] W.P. Lafuse, D. Brown, L. Castle, B.S. Zwilling, Cloning and characterization of a novel cDNA that is IFN- γ -induced in mouse peritoneal macrophages and encodes a putative GTP-binding protein, *J. Leukocyte Biol.* 57 (1995) 477–483.
- [40] D.A. Carlow, S.J. Teh, H.S. Teh, Specific antiviral activity demonstrated by TGTP, a member of a new family of interferon-induced GTPases, *J. Immunol.* 161 (1998) 2348–2355.
- [41] A.D. Luster, Chemokines—Chemotactic cytokines that mediate inflammation, *N. Engl. J. Med.* 338 (1998) 436–445.
- [42] P. Zhang, L.J. Quinton, G.J. Bagby, W.R. Summer, S. Nelson, Interferon- γ enhances the pulmonary CXC chemokine response to intratracheal lipopolysaccharide challenge, *J. Infect. Dis.* 187 (2003) 62–69.
- [43] E. Garcia Del Busto Cano, K.W. Renton, Modulation of hepatic cytochrome P450 during *Listeria monocytogenes* infection of the brain, *J. Pharm. Sci.* 92 (2003) 1860–1868.
- [44] H.F. Rosenberg, J.B. Domachowske, Eosinophils, eosinophil ribonucleases, and their role in host defense against respiratory virus pathogens, *J. Leukocyte Biol.* 70 (2001) 691–698.
- [45] T. Ozaki, S. Sakiyama, Molecular cloning of rat calpactin I heavy-chain cDNA whose expression is induced in v-src-transformed rat culture cell lines, *Oncogene* 8 (1993) 1707–1710.
- [46] J. van den Boom, M. Wolter, R. Kuick, D.E. Misek, A.S. Youkilis, D.S. Wechsler, et al., Characterization of gene expression profiles associated with glioma progression using oligonucleotide-based microarray analysis and real-time reverse transcription-polymerase chain reaction, *Am. J. Pathol.* 163 (2003) 1033–1043.
- [47] T. Nagata, Y. Takahashi, Y. Ishii, S. Asai, Y. Nishida, A. Murata, et al., Transcriptional profiling in hepatoblastomas using high-density oligonucleotide DNA array, *Cancer Genet. Cytogenet.* 145 (2003) 152–160.
- [48] S. Munoz, R. Hernandez-Pando, S.N. Abraham, J.A. Enciso, Mast cell activation by *Mycobacterium tuberculosis*: mediator release and role of CD48, *J. Immunol.* 170 (2003) 5590–5596.
- [49] E. Marshall, Getting the noise out of gene arrays, *Science* 306 (2004) 630–631.